R E P O R T R E S U M E S

ED 011 942 24 AA 000 059

PARTICIPANT TEACHER JUDGMENTS OF EXPERIMENTAL PROGRAMS IN SECONDARY MATHEMATICS.

BY- RYAN, JAMES J. RISING, GERALD R.
MINNESOTA STATE DEPT. OF EDUCATION, ST. PAUL

REPORT NUMBER BR-5-1028-TR-3 PUB DATE AUG 66
REPORT NUMBER TR-66-3
CONTRACT OEC-5-10-051
EDRS PRICE MF-$0.09 HC-$1.24 31P.

DESCRIPTORS- QUESTIONNAIRES, *INSTRUCTIONAL MATERIALS, *EXPERIMENTAL PROGRAMS, *TEACHER ATTITUDES, ITEM ANALYSIS, FACTOR ANALYSIS, *PROGRAM EFFECTIVENESS, COMPARATIVE ANALYSIS, GRADE 7, GRADE 8, GRADE 9, GRADE 10, GRADE 11, *MATHEMATICS INSTRUCTION, SAINT PAUL

THE JUDGMENTS OF TEACHERS AND THEIR REACTIONS TO THE EXPERIMENTAL MATHEMATICS PROGRAMS THEY WERE TEACHING IN CONNECTION WITH A PROJECT WHICH EVALUATED SEVERAL RECENTLY DEVELOPED PROGRAMS WERE OBTAINED AND ANALYZED. OBSERVATIONS AND JUDGMENTS ABOUT THESE PROGRAMS WERE OBTAINED THROUGH A QUESTIONNAIRE SENT TO TEACHERS WHO PARTICIPATED IN ONE OF THE SEVERAL EXPERIMENTAL PROGRAMS FOR THE 7TH THROUGH 11TH GRADES DURING THE 1964-1965 SCHOOL YEAR. THE SPECIALLY CONSTRUCTED QUESTIONNAIRE, WHICH IS ATTACHED TO THE REPORT, WAS DESIGNED TO ELICIT TEACHER REACTIONS TO THE COMPARATIVE EFFECTIVENESS OF THE EXPERIMENTAL PROGRAMS OVER THE CONVENTIONAL PROGRAMS IN RELATION TO THE FOLLOWING FACTORS--(1) INSTRUCTIONAL EFFECTIVENESS, (2) PREFERENCE FOR INSTRUCTIONAL USE, (3) INSTRUCTIONAL INPUT DIFFERENCES, (4) PUPIL, PARENT AND OTHER TEACHER EVALUATIONS, AND (5) JUDGMENTS CONSIDERING DIFFERENT LEVELS OF PUPIL ABILITY. THE TEACHERS' JUDGMENTS OF THE INSTRUCTIONAL EFFECTIVENESS OF THE EXPERIMENTAL PROGRAMS WERE GENERALLY QUITE POSITIVE, AND A MUCH MORE FAVORABLE RESPONSE WAS MADE FOR THE EXPERIMENTAL MATERIALS IN CONTRAST TO THE CONVENTIONAL MATERIALS. A HIGH PROPORTION OF THE TEACHERS QUESTIONED PERCEIVED HIGHER ABILITY STUDENTS AS RESPONDING MORE FAVORABLY TO THE EXPERIMENTAL MATERIALS AND THE LOWER ABILITY STUDENTS AS RESPONDING MORE FAVORABLY TO THE CONVENTIONAL MATERIALS. A RELATED REPORT IS AA 000 058. (GD)

Participant Teacher Judgments of Experimental Programs in Secondary Mathematics

by

James J. Ryan

Gerald R. Rising

Technical Report No. 66-3

Minnesota National Laboratory
Minnesota State Department of Education

August, 1966

Participant Teacher Judgements of Experimental Programs in Secondary Mathematics[1]

James J. Ryan and Gerald R. Rising

Examination of the qualities and characteristics of alternate instructional programs requires consideration of the judgements and reactions of teachers who have used the materials as well as an assessment of the effects of the materials in terms of pupil achievement.

Teacher judgements and reactions to the materials subsequent to their use for instructional purposes are important for several reasons. Of most significance is the fact that the teacher's reactions and judgements determine to a great extent how much the materials will be used subsequent to their tryout and under what conditions and for what purposes. Also teachers' attitudes toward the materials influence not only their own but the decisions of other teachers not having tried the materials. Such "experience" judgements may well be more influential in this regard than more objective statistical data concerning pupil achievement based upon a broader sample of pupils and conditions.

Secondly, teachers are in a position to observe certain qualities and characteristics of the materials not obtained or indicated readily, if at all, by other information such as scores pupils obtain on achievement tests. These observations may concern factors relevant to pupil performance or factors relevant to the use of the materials as instructional tools by teachers.

It also has to be recognized that teacher reactions to materials and the effects of the materials on pupil achievement are likely to be interdependent which consequently has methodological implications for assessment of the effects of the materials on pupil achievement. Comparisons among alternate instructional programs or materials concerned with differences in pupil achievement should take into account differences in teacher attitudes and evaluations of the materials by making comparisons between teachers that are similar in these respects.

This report is concerned with the judgements and reactions of teachers to the experimental mathematics programs they were teaching in connection with a project evaluating several recently developed programs in secondary mathematics. This activity was carried out as part of the Secondary Mathematics Evaluation Project which has as its primary objective the assessment of several experimental programs in secondary mathematics in terms of pupil achievement. The programs being studied were those developed under the auspices of: Ball State, Indiana State Teachers College; University of Illinois Committee on School Mathematics, (UICSM); The University of Maryland Mathematics Program, (UMMaP); School Mathematics Study Group, (SMSG). Observations and judgements about these programs were obtained via a

[1]This study was carried out in connection with the Secondary Mathematics Evaluation Project supported in part by funds granted by the Course Content Improvement Section of the National Science Foundation (grant No. G-25164) and in connection with Project 2747, Effects of Modern and Conventional Mathematics Curricula on Pupil Attitudes, Interests and Perception of Proficiency, supported in part under contract with the Bureau of Research, U. S. Office of Education (contract No. OE-5-10-051).

questionnaire sent to participant teachers having used one of the several experimental programs for the 7th - 11th grades during the 1964-65 school year. The purpose of the survey was to determine the nature of the participant teachers' judgements and attitudes concerning the experimental relative to the conventional or traditional programs in mathematics and to determine if there were any similarities in the attitudes and judgements of teachers following the same experimental programs. It was also of interest to determine whether there were similarities among teachers at the same grade levels and, because teachers had participated in the project for varying lengths of time (1 to 3 years), to determine whether the amount of teacher experience with the experimental programs influenced their attitudes or judgements.

A separate objective for which this data was gathered was to determine whether the teacher judgements or attitudes exhibit any correspondence with indices of pupil achievement and interest in mathematics and if so to consider this information in connection with the analysis of the effects for the experimental programs on pupil achievement and motivation. It should be noted however that this study was carried out to determine the trends and differences for this particular sample of participant teachers rather than to test certain a priori hypotheses which would have implications concerning these variables and conditions for teachers outside of this sample.

I. Method

A. Instrument

A questionnaire was constructed with items inquiring about factors and conditions relevant to the instructional use and outcomes of the experimental programs. The items for the most part dealt with teacher reaction and judgements in the following general areas:

a) Differences in content and instructional demands (i.e. instructional input differences) between the experimental (E) and the conventional (C) programs.

b) Instructional effectiveness of the E compared to the C programs.

c) Preference for subsequent instructional use of the E compared to other programs or materials.

d) Attitudes and reactions of pupils and others (parents and other teachers) to the E programs.

A large proportion of the items required the respondent to make an explicit or implicit comparison between the experimental and conventional programs they were teaching, e.g. "To what extent do the experimental materials require a different method of presentation of the content than you follow in the conventional class?" For most items, responses were indicated by selecting one of several alternatives or making a check at some point on a graphic scale. A copy of the questionnaire is included in the Appendix.

B. Sample

The questionnaire was sent to all teachers participating during the 1964-65 school year in the Secondary Mathematics Evaluation Project who were teaching one class following a conventional program of instruction in addition to one class, at the same grade level, following one of the experimental programs of instruction. Classes ranged from the seventh through the eleventh grade with the experimental classes following one of the two or three experimental programs at a given grade level as shown in Table 1. Teachers varied from one to three years with respect to the amount of experience they had had with the experimental programs.

The questionnaire was sent to a total of 129 teachers. Of this number 119 or 92 percent returned completed questionnaires which comprised the sample for purposes of analysis. The number of recipients and the number of responding teachers classified by grade, experimental program and number of years experience are shown in Tables 1 and 2.

Table 1

Teacher Attitude Questionnaire

Recipients by Grade, Years Experience and Program

Grade	Years Experience	Program				
		Ball State	UICSM	UMMaP	SMSG	Total
7	2, 3	0	-	2	4	6
	1	3	-	3	1	7
	Total	3	-	5	5	13
8	2, 3	10	-	1	4	15
	1	1	-	0	4	5
	Total	11	-	1	8	20
9	2, 3	9	7	-	7	23
	1	2	5	-	12	19
	Total	11	12	-	19	42
10	2, 3	3	2	-	7	12
	1	2	1	-	10	12
	Total	5	3	-	17	24
11	2, 3	7	-	-	9	16
	1	9	-	-	5	14
	Total	16	-	-	14	30

Table 2

Teacher Attitude Questionnaire

Respondents by Grade, Years Experience and Program

Grade	Years Experience	Program: Ball State	UICSM	UMMaP	SMSG	Total
7	2, 3	0	-	1	4	5
	1	2	-	3	1	6
	Total	2	-	4	5	11
8	2, 3	9	-	1	4	14
	1	1	-	0	4	5
	Total	10	-	1	8	19
9	2, 3	8	6	-	7	21
	1	2	4	-	11	17
	Total	10	10	-	18	38
10	2, 3	3	2	-	7	12
	1	1	1	-	7	9
	Total	4	3	-	14	21
11	2, 3	7	-	-	9	16
	1	9	-	-	5	14
	Total	16	-	-	14	30

C. Analysis

The analysis was carried out considering responses to individual items and scores on multiple-item indices. The multiple-item indices were derived by combining responses to separate items concerned with the general factors or conditions. The alternate response possibilities for the separate items were differentially weighted and the responses of an individual teacher summed across the subset of items to obtain a score on a given index. For example, an index representing the "instructional effectiveness of the E program" was obtained by combining the responses to three separate items each dealing with a more specific instructional effect. These scores were used to provide more reliable indices of general attitudes and judgements that appeared to be reflected by the individual items.

In addition to determining the nature of the distribution of individual item responses and multiple-item indices for all teachers, the analysis was directed toward determining the differences (a) among teachers following the alternate experimental programs within each grade level, (b) between teachers having more or less experience with the E program within and across grade levels, and (c) among teachers at the different grade levels. Either chi-square or an exact probability test was used in most instances to determine the reliability of these differences.

Since the purpose of the analysis was to describe the trends and differences in judgement exhibited by this particular teacher sample, many of the statistical

comparisons are of a post hoc nature. Therefore, unless the differences are quite highly reliable ($p < .01$) any interpretations or generalizations beyond this sample should be made with caution. Because of this only differences around the .05 level of probability or less will be given emphasis.

Also as can be seen in Table 2, there was not a balanced number of teachers among the program, grade and experience level categories and for some categories a small number of teachers. Consequently a systematic comparison for each of these conditions independent of (i.e. within levels of) the others was not always possible.

The differences observed by this analysis should then be considered more a guide for subsequent research, i.e. as suggestive of hypotheses to be tested with additional data, than as conclusive evidence of a more general difference.

II. Results

A. Perceived reactions of pupils, other teachers and parents to the E program.

1. Parent reaction

Two items asked the teacher to indicate the reactions of parents to the experimental programs being taught. One of these items asked about the amount of parent interest and inquiry concerning the E materials, the other about the nature of the parent reaction, i.e. whether positive or negative.

a. Amount of parent interest and inquiry.

For the total sample of teachers, responses ranged from "no interest" to "a great deal of interest" on the scale providing these two response alternatives at the extremes. The most frequent response was "some interest". Responses at the lower end of the scale, i.e. "some" or a lesser amount of interest accounted for 63 percent of the responses.

i. Program differences within grades

The only program difference observed was at the ninth grade level. A smaller proportion of the SMSG teachers reported a moderate or greater level of parent interest than teachers following alternate ninth grade programs (Ball State and Illinois). This difference which is statistically reliable at the .05 level of probability ($\chi^2 = 5.3$) is shown in Table 3.

Table 3

Reported level of parent interest
for ninth grade programs

Program	Low	High	
UICSM Ball State	11	9	20
SMSG	16	2	18
	27	11	38

ii. Experience differences

No differences in responses to this item were observed to be associated with the level of teacher experience either within or across grades or programs.

iii. Grade level differences

Teachers of 7th and 8th grade classes much more frequently indicated a greater amount of parent interest and inquiry than teachers of 9th - 11th grade classes. This difference which is statistically reliable ($\chi^2 = 16.8$, $p < .001$) is shown in Table 4.

Table 4

Number of teachers reporting high and low levels of parent interest for 7 - 8th and 9 - 11 grade levels.

Grade	"None" or "some" interest	"Moderate" or "great deal" of interest	
7 - 8	10	20	30
9 - 11	67	21	88
	77	41	118

This difference probably reflects a greater interest in pupil school work in general on the part of parents when pupils are in Junior High School than when they are in Senior High School. A similar difference has been noted for other parent participation activities such as P.T.A. It might also reflect the fact that there is a greater actual difference between the content of the experimental and conventional materials at these lower grade levels than at the higher grade levels in that the E programs provide an introduction to mathematics rather than a review of arithmetic. This would provide a more evident contrast between programs from the parents point of view and might elicit thereby more interest or concern. There were no other statistically reliable grade level differences.

b. Nature of parent reaction

With respect to the evaluative nature of the parents' reaction, it is only at the 7th and 8th grade levels that there was a sufficient indication of interest which could serve as the basis for a judgement in this regard. Of the 20 seventh and eighth grade teachers reporting a moderate or greater amount of parent interest, 16 indicated that the parent reaction was more positive than negative, and 2 that the reaction was more negative than positive, the remaining two giving qualified responses. Among the 21 teachers in the ninth to eleventh grades indicating a moderate or greater amount of parent interest, 9 characterized their reaction as more positive than negative, 6 as more negative than positive and the remainder gave neutral or qualified responses. The small proportion of teachers in the higher grade levels indicating any amount of actual parent interest precludes any meaningful comparisons between the grade levels. There was however no real indication of any degree of negative reaction at any grade level.

2. Teacher perception of pupil reaction to instructional materials.

Three separate items asked about the reactions of pupils in the teacher's experimental and conventional classes to the instructional materials they were using. To obtain an overall index for each teacher representing the degree to which the E and C class pupils were seen to respond more favorably to the E than the C materials, weights (from 1 to 3) were assigned to the alternate responses to each item according to the extent the response indicated a more positive evaluation of the E materials. The response weights were summed over the items to obtain an overall score or index for each teacher representing the degree to which they judged a more favorable pupil response to the E materials.

a. General index of pupil evaluation

Comparisons considering each condition (program, grade level and amount of teacher experience) indicated no statistically reliable differences among programs within grades, among grades or between levels of teacher experience for the derived index of pupil evaluation of instructional materials.

b. Individual item differences

i. General E relative to C class reaction.

Each of the separate items inquiring about pupil reaction and evaluation of the instructional materials was also examined for program, grade and teacher experience differences.

One item (#13) asked whether E pupils in general were enjoying mathematics more or less than the C class pupils. Over all grades and programs, 59 percent of the teachers indicated the E pupils were enjoying mathematics more than the C pupils, 28 percent indicated less and the remainder indicated no difference or did not respond. Considering only those giving an unequivocal response the difference between the proportion indicating E as more and less favorable (68 and 32 percent, respectively) and an expectation of equal proportions is significant at the .01 level. This overall difference resulted primarily however from a greater difference in this regard at the lower grade levels, i.e. grades 7 - 9 rather than at the higher grade levels as shown in Table 5.

Table 5

Number of teachers indicating E class pupils enjoyed math "more" and "less" or the same as C class pupils.

Grade	E more	E less or same	
7 - 8	21	9	30
9	24	14	38
7 - 9	45	23	68
10 - 11	23	25	48
	68	48	116[2]

Comparing the frequencies for those indicating E class pupils enjoyed mathematics more than C class pupils, the difference between the 7 - 9th grades and the 10 - 11th grades is fairly reliable statistically ($\chi^2 = 3.15$, $p < .10$) while that between the 7 - 8th grades and the 9th grade is not ($p \approx .50$). No program nor teacher experience within-grade differences were observed. There was then a general tendency for ... teachers in this sample to view E class pupils as responding more favorably than C class pupils. This tendency was somewhat stronger and more definite for 7 - 9th than 10 - 11th grade teachers.

ii. Reactions for pupils of different ability levels.

Another item (#14) asked teachers to make a similar comparison with respect to the "interest and enthusiam exhibited for mathematics" on the part of E and C class pupils. For this item, separate judgements were requested comparing E and C classes for high, average and low ability pupils respectively.

[2]The difference between the total number of teachers in a given table and the total number responding (119) is due to some teachers either not responding to a given item or providing a response which cannot be classified in the categories considered.

For all teachers, the responses to this item indicated a very definite relationship between level of pupil ability and the relative amount of interest pupils were seen as having for the E as compared to the C materials. This is shown in Table 6.

Table 6

Number of teachers indicating E class pupils as having more or less (and equal) interest in math than C class pupils for each pupil ability level.

Ability Level	E class more	C class more or the same
High	88	27
Average	50	57
Low	22	83

Considering all teachers, fifty-five percent indicated that for high ability pupils there was more interest on the part of those in the E class and that for low ability pupils, those in the C class had greater interest. Considering only the responses of teachers that did indicate an E - C difference in interest at each ability level (i.e. E > C or vice-versa rather than E = C), the tendency to make a differential judgement concerning relative interest for the separate ability levels is even more evident when the judgements for each ability level are cross-tabulated as shown in Table 7.

Table 7

Number of teachers making alternate judgements concerning relative interest of E and C class pupils for each ability level.

Judgement for low ability	Judgement for high ability: E > C	Judgement for high ability: C > E	
E > C	14	3	17
C > E	66	8	74
	80	11	91

There were no response differences in these judgements between programs within grades, grade levels, or levels of teacher experience indicating that this response tendency was independent of (i.e. not altered by) these conditions. There appears then to be a definite belief or judgement among teachers in this sample that high ability pupils in the E class develop more interest in mathematics than high ability pupils in the C class and that low ability pupils in the C class develop more interest than those in the E class, i.e. that E and C programs differentially affect pupils' reaction depending upon the pupil ability level.

iii. Judgement of pupils' relative preference for E and C materials.

Another item (#15) asked teachers to indicate for each class separately the relative extent to which pupils felt the E or C materials were better. For this item each teacher made a separate response for each class, consequently an indication of whether or not the E class pupils favored their materials more than the C class pupils favored theirs (or vice-versa), that is the relative preference of each class for its own materials, could be determined for each teacher. Of a total of 107 teachers responding to this item, 56 percent indicated a more positive reaction on the part of E pupils toward their materials than C pupils toward theirs, and 20 percent indicated a more positive reaction on the part of the C pupils than the E pupils toward their respective materials. The remaining teachers indicated the reaction of pupils in both classes to their own materials to be approximately the same. The difference between the proportions of teachers judging E pupils as being more favorable toward their materials than C pupils than the converse, 56 percent and 20 percent respectively, is highly reliable statistically ($p < .01$). Similar statistically reliable differences are observed at each separate grade level with the exception of the 10th grade where there was no difference in the proportions of teachers judging C pupils and those judging E pupils as favoring their respective materials. No within-grade program nor teacher experience differences were observed for this item.

Considering these three items concerned with the teachers' judgement of the pupils' reaction there was a definite tendency across all programs and grade levels, with the exception of the 10th grade, for teachers to report that E class pupils in general respond more favorably to their instructional materials than C class pupils to theirs. This perceived difference for the class in general occurred in spite of a fairly consistent perception of a more favorable reaction to the C materials on the part of the low ability pupils.

3. Perception of other teachers' evaluation of E programs.

Teachers were asked to indicate how other mathematics teachers in their school and grade felt about the experimental programs the participating teachers were using in terms of "the degree to which they favored or opposed the E programs".

Among the 94 teachers giving a definitive response to this item, approximately 38 percent indicated that there exists some degree of opposition among other teachers.

a. Within-grade program differences.

Only at the ninth grade level were any program differences apparent for this item. A smaller proportion of SMSG teachers reported a favorable response from other teachers than was the case for teachers following the other programs at the ninth grade level (Ball State and UICSM). This difference, which is shown in

Table 8 however, was not extreme enough to be highly reliable statistically, $\chi^2 = 3.44$, $.05 < p < .10$, when the Ball State and UICSM frequencies were combined for the comparison.

Table 8

Number of teachers following 9th grade programs indicating more and less favorable reactions to the E programs by other teachers.

Program	Less Favorable	More Favorable	
SMSG	9	6	15
Ball State	2) 4	7) 14	18
UICSM	2)	7)	
	13	20	33

b. Grade level and teacher experience differences

There were no reliable grade level or teacher experience differences for this item. However, it is interesting to note that at the ninth grade level the largest proportion of teachers reported a more favorable reaction to the E program by other teachers. It can be seen in Table 8 that this was primarily due to the responses of the teachers following the Ball State and Illinois programs who as a group reported with a considerable greater frequency favorable reactions by other teachers.

B. Judgement of amount of instructional input difference between the E and C programs

Five separate items in the questionnaire provided the basis for obtaining an indication of the extent to which teachers observe a difference in the instructional factors and characteristics (e.g. the content and presentation conditions) associated with the experimental and conventional materials or programs as they presented them. These items involved judgements concerning content differences, method of presentation, preparation effort and difficulty and amount of inservice training required for the E program they were teaching.

1. General instructional input difference index

To obtain a single E - C instructional difference index or score for each teacher, alternate responses to each item were assigned weights (from 1 - 3) according to the amount of difference between the E and C programs the response represented for a given item. The weighted responses were then summed to obtain an overall score for each teacher which was considered to represent the amount of difference in instructional input the teacher felt there was between the E and C programs as they were taught.

Program, grade and experience comparisons were made in terms of the number of teachers above and below the median of the distribution of "instructional difference" scores for all responding teachers.

a. Within grade program differences

The only statistically reliable within-grade program difference was observed at the 11th grade between the Ball State and the SMSG programs. The distribution of instructional-difference scores for these teachers is shown in Table 9.

Table 9

Number of instructional difference scores above and below the median for Ball State and SMSG teachers in the 11th grade.

Program	Below Median	Above Median	
Ball State	3	13	16
SMSG	10	4	14
	13	17	30

These results, which are highly reliable statistically ($p = .009$, Fishers' exact probability test), indicate that, at the 11th grade, teachers following the Ball State program much more frequently indicated differences in the conditions and activities relevant to the instruction they provide with their experimental materials compared to their conventional materials than did teachers following the SMSG program. This difference also appears to be independent of the amount of teacher experience with these programs. This suggests that from the teachers' point of view the Ball State program at the 11th grade differed to a greater extent from the usual conventional program than the SMSG program. However examination of the actual content of these two programs gives exactly the opposite impression, making this result quite unexpected and difficult to explain.

b. Experience differences

No reliable teacher experience differences within or across grades and/or programs were found for the index of instructional input differences.

c. Grade level differences

No reliable grade level differences were observed for this index.

2. Response to individual items concerned with instructional input differences.

Comparisons were also made for the separate items dealing with differences for specific kinds of instructionally relevant conditions which comprised the "instructional difference" index. Examination of responses to the individual items for the 11th grade teachers for whom a clear program difference was observed for the multiple-item index indicated that for each separate item response differences in the direction indicated for the overall index occurred though not to a statistically reliable extent in each instance.

a. Preparation effort

In response to the question concerning the relative ease of class preparation between the E and C classes, 81 percent of the responding teachers indicated less preparation difficulty for the C than for the E class. However only 28 percent indicated that C class preparation was "very much easier", the most extreme scale designation provided. That there was more preparation difficulty reported for the E class would be expected considering these were new and different materials. The more interesting question concerns whether there were any program or grade level difference in this regard. The appropriate comparisons indicated no within-grade program differences in response to this item.

With respect to grade level differences, the 7 and 8th grade teachers indicated much less frequently that the C class preparation was "very much easier" than that for the E class (in contrast to C being "somewhat easier" or E being easier) than did teachers in the higher grades ($\chi^2 = 5.10$, $p < .025$). This is shown in Table 10. The bigger differences in preparation effort reported by teachers for the 9th and higher grades may in part reflect a greater preparation effort in general for teachers at these grade levels which might be accentuated when new materials are used.

Table 10

Number of teachers at higher and lower grade levels indicating the extent C class preparation is easier than E.

Grade	E easier or C somewhat easier	C very much easier	
7 - 8	26	3	29
9 - 11	57	30	87
	83	33	116

At the 9th grade level there was also a statistically significant difference ($\chi^2 = 7.09$, $p < .01$) between levels of teacher experience, the less experienced teachers indicating much more frequently that C is "very much easier than E". This difference shown in Table 11 would seem to be expected although it did not appear at other grade levels. This result indicates that for the 9th grade programs at

least, preparation effort was related to experience with the experimental program.

Table 11

Number of 9th grade teachers at each level of experience indicating the extent C class preparation is easier than E.

Years experience	E easier or C somewhat easier	C very much easier	
1	6	11	17
2 - 3	16	3	19
	22	14	36

b. Difference in method of presentation

In response to the question concerning the degree of difference in the method of presentation of E and C class materials, approximately 18 percent of the responding sample indicated "a great difference" - the most extreme alternative provided and only 6 percent of the teachers gave responses indicating a slight or lesser difference at the other extreme. The greatest proportion of responses indicated some moderate degree of difference. There were no clear program, grade nor experience differences in response to this item.

c. Difficulty teaching for first time

In response to the question concerning the possible difficulty a teacher would experience following the E program for the first time, 92 percent of the teachers indicated at least "some difficulty" would be experienced and 25 percent indicated "considerable difficulty". There were nc clear grade, program or experience differences in response to this item.. The response to this item certainly suggests the need for some inservice instruction prior to use of experimental materials. Further evidence on this point is provided by the responses to the item (#4) asking for an indication of "the number of hours of inservice training" teachers should have prior to use of the experimental materials. Overall 21 percent of the teachers indicated more than 20 hours, 51 percent indicated more than 15 hours and approximately 75 percent indicated more than 10 hours of inservice instruction is needed prior to use of these programs. There were no clear program, grade or experience level differences for this item.

d. Difference in content

Responses to the item concerned with the degree of difference in content between the E and C class presentations indicated that approximately 58 percent of the teachers judged the difference to be quite large. There were no clear within-grade program differences or experience differences. There was however a definite

tendency for teachers at the lower grade levels to judge these differences as greater than did teachers in the higher grade levels. This is shown in Table 12. These differences appear to be quite in accord with the actual differences in content that exist between the experimental and typical conventional programs at each of these grade levels.

Table 12

Number of teachers indicating relatively larger and smaller differences in content between E and C classes at different grade levels.

Grade	Smaller difference	Larger difference
7 - 8	7	22
9	14	23
10 - 11	25	25

Comparing the responses to this item with those concerning differences in method of presentation, it appears that teachers saw as much if not more difference in the content than in the method. Among 7 and 8th grade teachers there were more frequent indications of content differences than method differences. For both the ninth and eleventh grades an approximately equal proportion, about one-half, of the teachers indicated a moderately large difference in content and methodology. At the 10th grade there was a smaller but equal proportion of teachers that indicated moderate or greater differences for each of these characteristics.

In general there was not a very high degree of correspondence between responses to items representing input differences between the E and C programs among teachers following the same program or at the same grade level. This seems to indicate that even though following the same E program, there is not alot of consistency or similarity among teachers with respect to the instructional input differences they perceive or respond to between the experimental and conventional programs. This also suggests that more extensive and specific inquiry concerning the nature of the E and C program differences is necessary to obtain a reliable indication of the ways in which the E programs are seen to differ from the conventional programs by the teachers.

C. Instructional effectiveness of the experimental programs.

Teacher judgements about the effects of the experimental compared to the conventional programs in terms of pupil achievement in mathematics were obtained from three separate items. These items asked the teacher to indicate whether pupils in the E or the C class (1) acquired the best background for future mathematics instruction, (2) exhibited the most understanding and comprehension of the material presented, and (3) acquired the most knowledge and proficiency with respect to the course content.

1. General index of instructional effectiveness

A single index representing the teachers' overall judgement of relative effectiveness was obtained by considering the frequency (0, 1, 2, or 3) over the three separate items with which the teacher responded by indicating that the E class rather than the C class was more effective. Comparisons were then made in terms of the proportion of teachers in the three frequency categories. These questions requested judgements which would reflect the performance or achievement of the pupils in the specific E and C classes. It was recognized that in some cases a performance difference could be the result of actual differences in initial ability between the pupils in the two classes in addition to or rather than differences resulting from the instructional materials. To have some check on this possibility and identify such cases, in connection with two of the questions concerning the effectiveness of the E program, an additional question was asked. This question concerned whether the difference noted between the classes was primarily due to differences in ability or to the instructional materials.

Overall 28 percent of the responding teachers indicated in connection with one or both items that initial ability differences accounted for or were the primary source of the judged difference in relative effectiveness between the two classes. Because of this these teachers were not considered further in the analyses concerning the judged effectiveness of the programs.

Considering all remaining teachers, 46 percent responded by indicating that the E class had achieved all three of the objectives to a greater extent than the C class; an additional 21 percent indicated E for two of the three items. Twenty-four percent did not indicate that the E class had achieved any of these objectives to a greater extent than the C class. (It should be noted here that not responding in favor of E is not always the same as responding in favor of C since a small proportion of respondents indicated equivalence between E and C on one or another of the items).

a. Within-grade program differences.

The only discernible program differences were observed at the 9th grade level. There was a tendency for the SMSG teachers to judge the E program as more effective than the C program less frequently than teachers of the other 9th grade programs (Ball State and UICSM). This difference which is shown in Table 13 exhibits only a moderate level of statistical significance, $p = .05$, Fisher's exact probability test.

Table 13

Responses of ninth grade teachers indicating the relative effectiveness of the experimental program they taught.

Program	Number of objectives for which E indicated as more effective than C program: 0 - 1	2 - 3	Total
Ball State	3	5	8
UICSM	1	7	8
	4	12	16
SMSG	6	3	9
	10	15	25

b. Between grade differences

As shown in Table 14, the 11th grade had the largest proportion of teachers giving the most positive evaluation of the E program and the 9th grade had the smallest proportion. Separate comparisons between the distributions for the 9th and the 11th grade teachers, respectively, and those in the remaining grades (7, 8 and 10) indicate that the difference for the 11th grade was quite reliable ($\chi^2 = 5.92$, $p < .025$) but that for the 9th grade was not.

Table 14

Number of teachers at various grade levels indicating more and less frequently that the E class has achieved designated objectives to a greater extent than the C class.

Grade	Number of objectives[3] for which E indicated as better than C: 0 - 2	3	
7 - 8	12	12	24
9	20	5	25
10	8	4	12
11	3	15	18
	43	36	79

[3]The comparison was made in terms of 0-2 vs. 3 responses favoring E because this division gave an approximately equal proportion of cases in each response category.

It appears that teachers following the 11th grade programs judged their experimental programs to be more effective than their conventional programs with respect to the instructional objectives considered with greater frequency than did teachers at other grade levels.

c. Teacher experience

When responses to all three items were considered, no differences associated with levels of teacher experience were observed within or across grades or programs.

2. Individual item differences.

The trends noted above for the general index of effectiveness tended to be observed for each of the separate items comprising the index. With respect to the program differences at the ninth grade level the SMSG teachers significantly less frequently ($p < .05$ Fisher's exact probability test) than the UICSM and Ball State teachers indicated that the E class acquired (1) the best background for future mathematics and (2) the most knowledge and proficiency. With respect to the latter objective none of the SMSG teachers indicated the E class as being superior to the C class. These results are shown in Table 15. The differences observed for the SMSG teachers for these two items accounts for the difference observed for the SMSG teachers for the overall index of effectiveness noted above.

Table 15

Responses of ninth grade teachers indicating judgements of relative effectiveness of E programs with respect to specific objectives.

Program	Class acquiring best background for future mathematics: E	Not E	Class acquiring most knowledge and proficiency in algebra: E	Not E
Ball State	6) 14	2) 2	4) 7	4) 9
UICSM	8)	0)	3)	5)
SMSG	3	6	0	9

There were also some differences between the items in the overall proportions of teachers indicating "E" was more effective with respect to a given objective than C which are shown overall and for each grade level in Table 16. Seventy-three percent of the teachers responded in favor of E to the item concerned with the "best background for future mathematics instruction"; 60 percent responding with "E" to the item concerning "understanding and comprehension of the materials;" and 52 percent indicated "E" for the item concerning "acquisition of knowledge and proficiency." There was then a greater tendency for teachers to judge the E program as relatively more effective with respect to providing a "better background for future mathematics" than to judge E as more effective with respect to providing the "most knowledge and proficiency". As can be seen in Table 16, this differential response occurred primarily for programs at the 7 - 10th grade levels, the 11th grade teachers responding in favor of the E programs to the same extent for all of the specific objectives. It is quite likely that in response to these items, teachers are

considering "knowledge and proficiency" as representing the development of computational skills.

Table 16

The number of teachers at each grade level indicating whether the E or C class was more successful in achieving each of the designated instructional objectives.

Objective:	Best background for future mathematics			Understanding and comprehension of material			Acquisition of knowledge and proficiency in subject		
Grade	E	C	No differ.	E	C	No differ.	E	C	No differ.
7 - 8	19	3	2	14	9	1	16	5	3
9	17	5	3	15	9	1	7	16	2
10	7	4	1	3	8	1	3	7	2
11	15	3	0	15	3	0	15	3	0
	58	15	6	47	29	3	41	31	7

3. Judgements concerning pupils for whom E programs are more appropriate.

Among other judgemental criteria concerning the effectiveness of instructional materials is the question of whether or not the materials might be adequate or appropriate for some pupils rather than others. To determine, therefore, whether teachers had reservations about the broad general use of the experimental programs they were following, they were asked to indicate "for what type of pupil or school" they felt the program would be "most appropriate". This question was asked in open-end form providing no a priori alternatives to choose among.

Sixty percent of the teachers responded to this question with the designation of "above average" pupils. An additional 13 percent gave responses which referred primarily to higher ability pupils, e.g. "college prep". Seventeen percent indicated "average and above". Only 8 percent gave responses that did not distinguish in some way with respect to ability and in the direction of favoring the E programs for higher ability pupils (e.g. "all"). There were no respondents indicating "below average" or simply "average pupils" in response to this question.

There were no within-grade program, grade or experience differences in response to this item.

The great majority of these teachers apparently feel the experimental programs are less appropriate for the lower ability or non-college bound pupil. Since these programs for the most part were developed primarily for college-bound pupils, this strong response tendency no doubt reflects in part an awareness of this fact. It is not clear however from the responses given whether the specification of "higher ability pupils" refers to those of relatively higher ability that are enrolled in the mathematics classes or to the above average pupil in general which would include a large proportion of those enrolled in mathematics classes from the 9th grade on. The response to this item also does not indicate how strongly teachers feel about this nor how great a difference between higher and lower ability pupils they believe there is in the appropriateness of the programs.

The responses to this item are very much in line with the teachers' perceptions that higher ability pupils respond more favorably to the E programs and lower ability pupils more favorably to the C programs. The data obtained in this survey provide no way of determining whether this perception influences the judgement of appropriateness or vice-versa or whether both stem from a more general belief or judgement about the nature or characteristics of the instructional materials. It is also a separate question as to whether these judgements and perceptions are in fact valid.

D. Teacher preference for use of the experimental materials.

Several items (#17 and #18) were directed toward obtaining an indication of the teacher preference for use in his own classes of the experimental program materials in contrast to conventional materials. An additional open-end item (#19) invited the respondent to give any additional comment he cared to make concerning the instructional materials. Alternate responses to these items were assigned weights in the direction of the degree of favorability they reflected toward the E program to obtain a single score index representing the degree of preference each teacher had for the E program he was teaching.

1. Index of preference for E program

Although for most programs and at all grade levels there was considerable variation in the scores obtained, comparisons among alternate programs within grades, by grade level and by level of teacher experience revealed no statistically reliable differences associated with these factors. At the 10th grade level there was a non-significant tendency for teachers to have lower preference index scores compared to teachers in other grades.

2. Individual items

In response to the item asking teachers to indicate their relative preference for use of E as compared to C materials for the next year, 67 percent indicated a preference for their experimental materials. There were no program, grade or experience differences in response to this item. In response to a separate item asking which materials, E, C or other, they would select to use if certain practical considerations (e.g. others' judgements, etc.) were not involved, 32 percent indicated their present "E materials in all classes"; an additional 11 percent indicated E materials in certain classes; 17 percent indicated their present C materials in all classes; and 35 percent indicated they would choose different materials (i.e. neither the same E nor C materials), the majority of the latter not providing any further characterization of their choice. Here again no differences were observed among alternate programs, grades or levels of teacher experience. Responses to these items seem to indicate that although there is somewhat more satisfaction with the E materials they were using than their present conventional materials there is a reasonably large proportion of teachers who would prefer using materials different from either.

E. Correspondence between separate judgement indices

To obtain an indication, in a more general way, of the extent to which the different indices of teacher reactions, judgements or perceptions were reflecting relatively separate and independent judgements about the qualities and characteristics of the E programs, the degree of correspondence between the separate indices of judged program effectiveness, of preference for the E program, of perceived pupil

evaluation and of instructional input difference were determined. Moderately high and statistically reliable ($p < .005$) relationships were found between the index of judged effectiveness, the index of teacher preference for the E program and the index of perceived pupil evaluation. There was no correspondence however between the latter three indices (which were of a more evaluative nature) and the index representing the degree of difference in instructional input characteristics between the E and C programs.

A further analysis was made to determine the degree of association between the teachers' index of preference for the E program and their separate characterization of parents' and of other teachers' reactions to the E programs. A tendency was found for the index of preference to be associated with both of these judgements especially the characterization of the other teachers' reactions.

The fact that there is a good deal of correspondence between the teachers' judgement of effectiveness of the E program, their preference for the program and their characterization of the evaluative reaction of others to these programs indicates that there may be a more general evaluative reaction factor common to and possibly influencing the teachers' response in each of these separate judgement areas.

Although the high correspondence observed between the teachers' judgement of effectiveness and their preference for instructional use of the E program is reasonable and to be expected, the correspondence between these indices and judgement of the reactions of others suggests that the latter may be more a reflection of the teacher's own reaction than an objective characterization of an independent source of evaluation. This interpretation seems more likely than the possibly alternate explanations simply on the logical grounds that teachers are in a better position to make objective judgements about performance and achievement effects than about pupils' and others' evaluative reactions to the materials. However, additional data, not collected as part of this survey is necessary to answer this question.

Summary and Discussion

There are several more general observations that can be made concerning the reactions and judgements of participating teachers to the E programs they were individually teaching which are indicated by this survey.

It should be recognized that the observations for the sample of teachers as a whole have to be interpreted with some caution since there is no base against which to compare the evaluations and judgements to determine the extent they reflect more objective elements in the teachers' experience in contrast to subjective factors such as "participation" effect (i.e. the "Hawthorne effect") and/or a priori expectations concerning the characteristics of the "new" programs. Differences resulting from comparisons made within the responding sample between different subgroups of teachers (representing alternate programs or grades) can be interpreted more meaningfully however because of their relative nature.

Judgement of instructional effectiveness.

The teachers' judgements of the instructional effectiveness of the E programs were generally quite positive in that they were judged to achieve certain designated objectives much more frequently than the conventional programs the teachers were using.

There tended to be some differences however in judgements of relative effectiveness among the alternate objectives. A larger proportion of teachers indicated the E program as being relatively more effective than the C program for providing a "background for future mathematics instruction" than for the facilitation of "knowledge and proficiency" in the specific mathematics subject. This possibly reflects a distinction between the development of skills specific to a given area of mathematics instruction and the development of a conceptual understanding of the general structure of mathematics. If so, this difference may indicate somewhat greater reservations concerning the E program with respect to development of more specific mathematical skills.

There were unexpectedly no reliable differences between teachers following different programs with respect to judgements of effectiveness.

There was, however, a clear tendency for the 11th grade teachers to consistently judge their E programs (Ball State or SMSG) as more effective for achieving the instructional objectives than teachers at other grade levels. This may reflect certain characteristics of the traditional or conventional program against which the E program is compared at this grade level as much as it reflects differences in the quality of the E programs for the 11th grade compared to other grade levels.

Preferences for instructional use.

Considering the teachers' preferences for materials for their own instructional purposes, there was, as observed for judgements of instructional effectiveness, a much more favorable response (i.e. preference) for the E materials in contrast to the conventional materials the individual teachers were using. However, when given an opportunity to indicate preferences for instructional materials among a broader range of alternatives, more than a third of the teachers chose materials other than the E or C materials they had been using. These judgements were representative of the entire sample of teachers, there being no clearcut program or grade differences in this regard and suggest that none of the materials considered are completely satisfactory for a fair sized segment of this sample of teachers.

Judgements concerning instructional input differences.

One definite program difference was observed with respect to the teachers' characterization of the instructional input differences between their E and C classes. At the 11th grade level, the Ball State teachers much more frequently indicated differences between the conditions and activities connected with the instructional input required for their E and C classes than did teachers using the SMSG program. Although this difference could be reflecting differences in the conventional program against which these two groups of teachers were comparing their E program or initial differences between the two groups of teachers choosing to follow these alternate programs, it is more likely reflecting some real differences in the nature of the two programs. It is significant to note that teachers using each of these programs did not differ in terms of the high proportions giving positive judgements of the relative effectiveness of the E programs.

In connection with instructional differences, some fairly clear indications were obtained of the need for inservice or other teacher instruction in conjunction with their initial use of the experimental programs. Indications of this need were quite consistent across all programs and grade levels.

Pupil, parent and other teacher evaluations.

The teacher judgements or perceptions of pupil, parent and other mathematics teacher evaluations of the E programs were found to correspond quite closely to their own judgements about the programs suggesting that the teachers' judgements of the evaluative reactions of others may have been influenced to a certain extent by their own judgements. This interpretation seems quite reasonable considering the usually subjective nature of such judgements with respect to pupil and parent reactions (unless the reactions are quite strong), and the relatively more objectively based judgements that teachers can make about the effectiveness of the programs.

The reported amount of parent reaction and interest was quite limited and appears to occur mainly in the 7th and 8th grades. Whatever the basis for the judgement, there also appeared to be no evidence of any substantial degree of negative parent reaction.

It is clear, however, that evidence other than that obtained in this survey is needed to determine the extent and nature of the reactions of pupils, parents and other teachers concerning the E programs and whether these judgements are influencing those of the teacher or vice-versa.

Teacher judgements of the E program considering different levels of pupil ability.

Probably the most striking and certainly most unequivocal judgement or perception reflected in the teacher responses is the differential reaction concerning the materials with respect to different levels of pupil ability. This is indicated by the high proportion of teachers who perceive higher ability pupils as responding more favorably to the E materials and lower ability pupils more favorably to the C materials and the high proportion who judge the E materials as most appropriate for higher ability or college bound pupils.

The objective basis for this judgement or belief may be somewhat open to question, however, in part because of the evidence indicated above concerning the possible subjectivity of judgements concerning pupil evaluations. Whatever the basis it is quite clear that there is a definite belief independent of grade or program, that the experimental programs, which are prototypes of the current curriculum revision in secondary mathematics, are most appropriately used for the instruction of higher ability or college bound pupils.

In general, because of the limited number of teachers represented for some of the programs and the technical imperfections of the questionnaire itself as well as other methodological limitations, only limited generalizations of the results obtained from this questionnaire survey to a larger population are probably warranted. Rather the results should be considered indicative of certain tendencies or trends which should be the focus of more extensive examination in subsequent studies.

Appendix:

A. Mathematics Teacher Questionnaire

Mathematics Teacher Questionnaire

Instructions:

The following questionnaire concerns comparisons between the two mathematics classes you are teaching which are participating in the secondary mathematics evaluation project this year. We are interested in obtaining your judgements concerning the pupil reactions to the materials used in these classes and your judgements about the materials themselves.

Read each item carefully. Indicate the answer that best represents your own judgement or opinion either by checking one of the alternatives given or by checking a point along the scale provided when one of these types of response is required. Please answer each question. If you would care to provide a comment in addition to your answer, feel free to do so. If on a given question, none of the alternatives seems to provide an appropriate answer, please answer in your own terms.

In the questionnaire E refers to the experimental class, the class using the experimental materials, and C refers to the conventional class at the same grade level.

1. How do the two classes (E and C) compare with respect to the effort required for your own preparation this year?

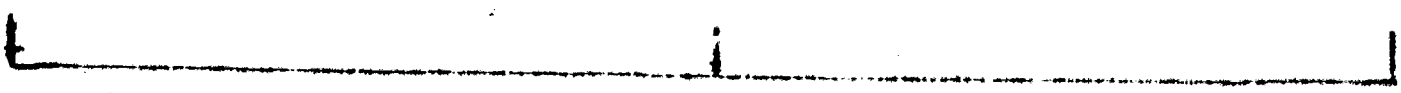
E class very much easier	E class somewhat easier	C class somewhat easier	C class very much easier

Are the differences indicated here due to content differences _____ or to pupil differences _____ between the two classes? (check one)

2. To what extent do the experimental materials require a different method of presentation of the content than you follow in the conventional class?

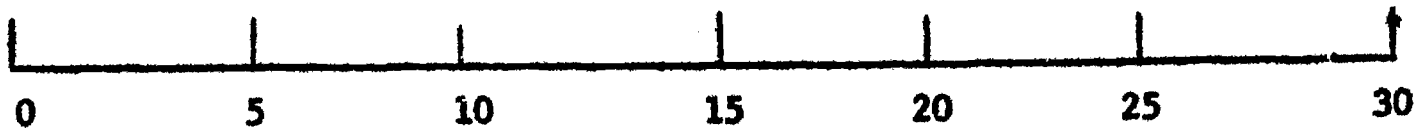
No difference	Some difference	A great difference

3. To what extent do you believe the average mathematics teacher (trained using conventional materials) would have difficulty teaching the experimental materials for the first time?

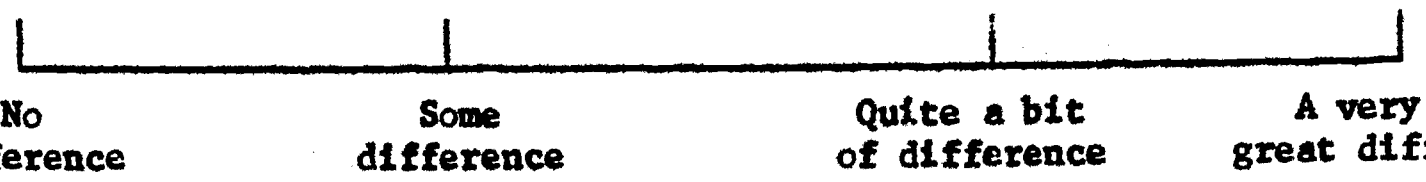


Considerable difficulty	Some difficulty	No difficulty

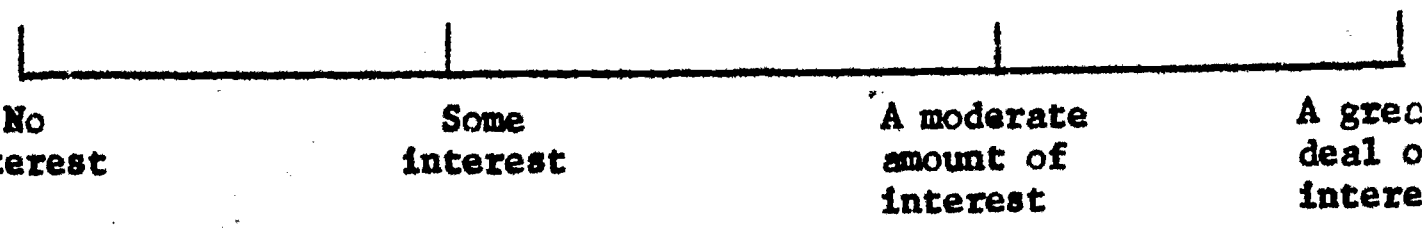
4. Approximately, how many <u>hours</u> <u>in-service</u> <u>training</u> do you believe the average mathematics teacher, trained in use of conventional materials, should have to teach the experimental materials you are using this year?



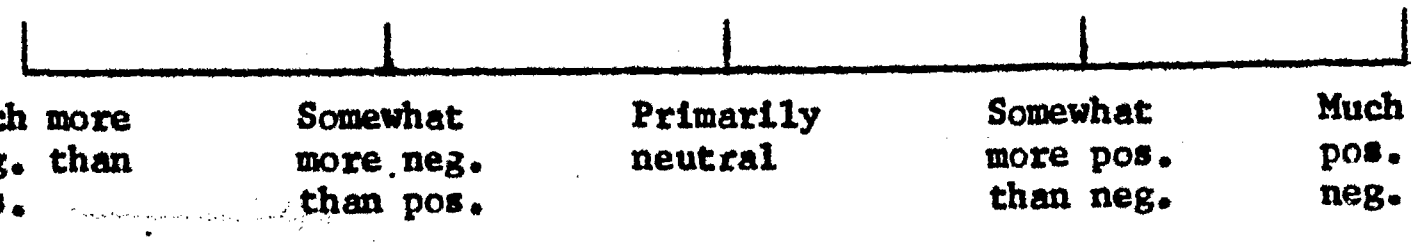
5. How much difference in general do you feel there is between the content of the experimental class materials and the content of the conventional class materials as you have introduced them?



6. How much interest or inquiry about the experimental materials have you received from parents of pupils in the experimental class?



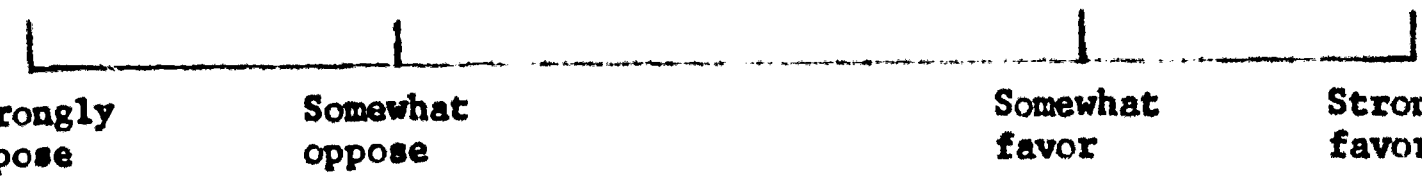
How positive or negative has the response from the parents been?



7. For what types of schools or types of students do you believe the experimental program you are teaching would be especially appropriate? _____

_____

8. In general, what do you believe would be the feeling of other secondary mathematics teachers in your school concerning the introduction of the experimental program in their classes at the same grade level as your class?



9. To what extent are the same tests used in each of the two classes? (check one)

_____ Both classes get exactly the same tests throughout the year.

_____ Both classes get the same major examinations such as midyear and final exams but different tests otherwise.

_____ None of the tests are exactly the same but a certain proportion of the questions are the same.

_____ The two classes receive two entirely different sets of tests.

_____ Other (elaborate) _____

_____

10. In which of the two classes do you believe the pupils are acquiring the best background for future mathematics instruction? E _____ C _____

11. In which class have the pupils usually exhibited more understanding and comprehension of the material presented? E _____ C _____

Are the differences indicated here primarily due to differences in pupil ability _____ or differences in instructional materials _____? (check one)

12. In which of the two classes do you believe the pupils are acquiring the most knowledge and proficiency in algebra? E _____ C _____

Are the differences indicated here primarily due to differences in pupil ability _____ or differences in instructional materials _____? (check one)

13. Do you think the pupils in the experimental class are enjoying mathematics more or less than pupils in the usual conventional classes you have taught?

more _____ less _____

14. In which of the two classes has there been a more favorable response to the materials in general (i.e. more interest and enthusiasm) on the part of each of the following groups of pupils:

	Experimental	Conventional
Higher ability pupils	_____	_____
Average ability pupils	_____	_____
Lower ability pupils	_____	_____

15. The pupils in the two classes are no doubt aware that their instructional materials differ. Either from what they may have heard or observed, they probably have formed some judgement about the different materials. On the following scales indicate how you think the pupils in each of the two classes feel about the materials in their class relative to the materials used by the other class.

Experimental class:

C materials much better | C materials somewhat better | E materials somewhat better | E materials much better

Conventional class:

16. To what extent do the pupils in the experimental class feel their materials are easier or more difficult than the materials used by the conventional class?

17. What is your own personal preference with respect to teaching either the experimental or conventional materials in your classes next year?

18. Independent of the judgement of others, and the actual availability of the materials, indicate which of the following best represents your present feeling concerning the materials you would decide to use in your ninth grade algebra class next year. (Your response to this item in no way commits you nor will be interpreted as bearing upon your participation in the experimental program next year)

_____ Use the present experimental materials in all algebra classes next year.

_____ Use the present conventional materials in all algebra classes next year.

_____ Use an entirely different set of materials next year. What materials?

_____

_____ Other (be specific) _____

_____

19. Use this space to provide any additional comments you care to make in relation to these items or about the experimental materials or the evaluation project.